

5307:10

ORC 66-8
March 1966

AD632205

BACTERIAL EXTINCTION TIME AS AN EXTREME VALUE PHENOMENON

by

Benjamin Epstein

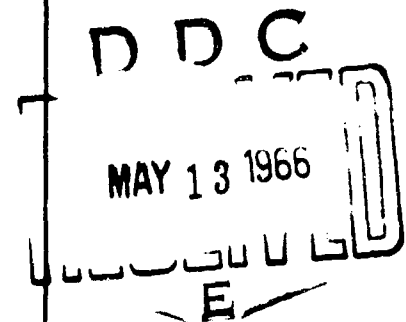
CLEARINGHOUSE FOR FEDERAL SCIENTIFIC AND TECHNICAL INFORMATION			
Hardcopy	Microfiche	9	pp
\$1.10	\$.50		
ARCHIVE COPY			

Code 1

OPERATIONS RESEARCH CENTER

COLLEGE OF ENGINEERING

UNIVERSITY OF CALIFORNIA - BERKELEY



BACTERIAL EXTINCTION TIME AS AN EXTREME VALUE PHENOMENON

by

Benjamin Epstein

Consultant, Palo Alto, California
and
University of California, Berkeley

March 1966

ORC 66-8

This research has been supported by the Army Research Office under Contract DA-31-124-ARO-D-331, and the National Institutes of Health under Grant GM-9606 and the National Science Foundation under Grant GP-4593 with the University of California. Reproduction in whole or in part is permitted for any purpose of the United States Government.

SUMMARY

The theory of extreme values is shown to be relevant in the analysis of two stochastic biological phenomena.

"Bacterial Extinction Time as an Extreme Value Phenomenon"

by

Benjamin Epstein

In [3], Mather found that some data on bacterial extinction times in the presence of a bactericide appeared to fall along a straight line when one plotted $\log(-\log p_t)$ against t , where p_t is the observed proportion of samples with no surviving bacteria after an exposure of t minutes to the bactericide.

In Mather's data, p_t was observed in samples of size 18 at times $t = 12(2)26$ minutes, and p_t varied from zero at $t = 12$ to $\frac{17}{18}$ at $t = 26$. Mather was led by heuristic reasoning to the model $\log(-\log p_t) = \alpha + \beta t$, where p_t is the "true" proportion of samples showing the absence of bacteria after an exposure of time t to the bactericide. Subsequently, it was noted by Gumbel on pages 43-44 of [3] that Mather had given an interesting application of the theory of extreme values without, however, using the term, since $p_t = e^{-e^{(\alpha + \beta t)}}$

is the type 1 asymptotic distribution of largest values. Gumbel then gave an alternative method of fitting the experimental data, which he deemed somewhat more appropriate in the light of extreme value theory. Actually, there appears to be little to choose between the two methods of fitting the data, since, in this instance, the parameter estimates obtained by either method are in excellent agreement with each other.

Neither Gumbel nor Mather gave a theoretical model for the time to extinction problem. To provide such a model would seem to be of some interest, particularly in view of the last paragraph on page 44 or [2]: "The three figures show that the theory of largest values is an adequate tool for the analysis of the extinction time of bacteria subject to a disinfectant. However, the philosophical question of why these data follow the law of largest values instead of smallest, as one might expect, remains open."

We now give a simple model for the bacterial extinction time problem which leads in a natural way to the type 1 distribution of largest values. In the model we make the following assumptions:

- a) the number of bacteria in a sample just prior to the introduction of a bactericide at time $t = 0$ is a Poisson random variable[†] with mean γ_0 ;
- b) the "lifetime" of a bacterium in the presence of the bactericide (a fixed dose is assumed) is distributed with c.d.f $F(t)$, $t \geq 0$.

It is clear that the presence or absence of bacteria in the sample after a length of time t has elapsed depends upon whether or not all of the bacteria, x , present in the sample at time $t = 0$ have "died" or, putting it another way, whether or not the longest lived bacterium survives. Since

$$p(x) = \frac{e^{-\gamma_0} \gamma_0^x}{x!}$$

is the probability that x bacteria are in the sample at time $t = 0$ and $[F(t)]^x$ is the probability that all x bacteria "die" by time t (assuming that bacterial extinction times are mutually independent) it follows from the theorem of total probability that $G(t)$, the c.d.f of the time T to total bacterial extinction in the sample, is given by^{††}

$$\begin{aligned} G(t) = \Pr(T \leq t) &= \sum_{x=0}^{\infty} \frac{e^{-\gamma_0} \gamma_0^x}{x!} [F(t)]^x \\ (1) \qquad &= e^{-\gamma_0 [1 - F(t)]}, \quad t > 0. \end{aligned}$$

[†] If the sample consists of v cc drawn from a solution in which bacteria are distributed at random with δ bacteria per unit volume, then $\gamma_0 = \delta v$.

^{††} A simpler proof of (1) is as follows: If x_0 , the number of bacteria in the sample at time $t = 0$, is Poisson distributed with mean λ_0 , the number of bacteria living longer than time t , is Poisson distributed with mean $\gamma_0 [1 - F(t)]$. Consequently, the probability that the sample will be bacteria-free is $\Pr(x_t = 0) = e^{-\gamma_0 [1 - F(t)]}$. But the events $T \leq t$ and $x_t = 0$ are equivalent and this establishes (1).

A discrete amount of probability $e^{-\gamma_0}$, which corresponds to the probability that there are no bacteria in the sample at time $t = 0$, is assigned to the c.d.f $G(t)$ at $t = 0$, i.e., $G(0^+) = e^{-\gamma_0}$.

If bacterium lives are distributed with c.d.f $F(t) = 1 - e^{-\lambda t}$, $t \geq 0$, an assumption that would appear to be reasonable in this situation, $G(t)$ becomes

$$(2) \quad G(t) = e^{-\gamma_0 e^{-\lambda t}}, \quad t \geq 0.$$

The c.d.f (2) is called the type 1 distribution of largest values in Gumbel's terminology. It follows from (2) that

$$(3) \quad \log(-\log G(t)) = \log \gamma_0 - \lambda t.$$

This is a theoretical justification of the model used to fit the data in [3].

A useful measure of the effectiveness of the bactericide is \tilde{t} , the modal or most probable time at which bacterial extinction takes place. \tilde{t} is the solution of the equation $g'(t) = 0$, where $g(t)$, the p.d.f of the time to extinction, is given by

$$(4) \quad g(t) = G'(t) = \gamma_0 \lambda e^{-\lambda t} e^{-\gamma_0 e^{-\lambda t}}, \quad t > 0.$$

It is readily verified that the maximum value of $g(t)$ is attained at

$$(5) \quad \tilde{t} = \frac{\log \gamma_0}{\lambda}.$$

The expected number of bacteria surviving in a sample after exposure to the bactericide for a length of time \tilde{t} is $\gamma_0 e^{-\lambda \tilde{t}} = 1$. Similarly, $G(\tilde{t}) = e^{-1}$. Thus, \tilde{t} is that exposure time such that, on the average, one bacterium in the sample still survives and it is also that exposure time for which 37%[†] of the samples become bacteria free. As pointed out by Mather, \tilde{t} is a simpler measure of the time to extinction in a sample than the more

[†] This is not to be confused with the probability of death of a single bacterium in the presence of the bactericide. If the time to extinction distribution of individual bacteria is given by $F(t) = 1 - e^{-\lambda t}$, then $F(\tilde{t}) = 1 - \frac{1}{\gamma_0} = \frac{\gamma_0 - 1}{\gamma_0}$. This the probability that an individual bacterium will die on or before time \tilde{t} .

common $t_{.50}$, the time at which 50% of the samples become bacteria free. $t_{.50}$ is related to \tilde{t} by the formula

$$(6) \quad t_{.50} = \tilde{t} - \frac{\log \log 2}{\lambda}$$

as can easily be seen from the relations $e^{-\gamma_0 e^{-\lambda t_{.50}}} = 1/2$. More generally, t_p , the time at which 100p% of the samples become bacteria free, is related to \tilde{t} by the formula

$$(7) \quad t_p = \tilde{t} - \frac{\log \log \frac{1}{p}}{\lambda}$$

It is also useful to point out that if the sample drawn from the solution to which bactericide had been added contained kv c.c. rather than v c.c., then the modal extinction time \tilde{t}_k for samples of this size would be

$$(8) \quad \tilde{t}_k = \frac{\log (k\gamma_0)}{\lambda} = \tilde{t} + \frac{\log k}{\lambda}$$

Our discussion has been purely theoretical under the assumptions that the model resulting in equation (2) is a reasonable description of the bacteriological situation. In practice, of course, γ_0 and λ may be unknown and would be estimated either graphically or analytically by the methods described by Mather and Gumbel. The value of the model is that it assigns a meaning to each of the parameters. For the data in [3], γ_0 and λ can be estimated respectively (see pages 140-141 of [3]) as $\gamma^* = 794$ and $\lambda^* = .376$. Thus we can say from the data that the number of bacteria in the sample at time $t = 0$ is Poisson distributed with estimated mean = 794 and that in the presence of 1.15% Phenol the life of an individual bacterium is exponentially distributed with estimated mean life $\frac{1}{\lambda^*} = 2.66$ minutes. Substituting in (5), (6) and (7), one obtains 17.75, 18.73, and 25.65 minutes respectively as estimates of \tilde{t} , $t_{.50}$ and $t_{.95}$. Substituting in (8), one obtains $17.75 + 6.12 = 23.87$ minutes as an estimate of

the modal extinction time in a 10v cc sample.

Comparisons between the effectiveness of the same bactericide at various dosages or between different bactericides at the same dosage, can be made by comparing the corresponding λ values.

The model for bacterial extinction times given in this paper leads to the type 1 distribution of largest values. It is interesting to note that the model given recently by Gart [1] for the distribution of response times after inoculating a host with a solution containing some microorganisms can also be considered in the framework of extreme value theory, the only difference being that response times are distributed as smallest values. The key points in the model are: 1) the number of inoculating particles in the sample is a Poisson random variable with mean γ , 2) the incubation time for a single particle is assumed to be a random variable with c.d.f $F(t)$, 3) a positive response occurs on or before time t if and only if at least one of the x inoculating particles present at time $t = 0$ has incubated, or putting it another way, if and only if the particle taking the shortest time to incubate has done so by time t . It follows readily from these assumptions that the c.d.f of the response time is $H(t) = 1 - e^{-\gamma F(t)}$, $t \geq 0$. If one lets the incubation time c.d.f for a single particle be $F(t) = [1 - e^{-\lambda t}]^N$ as is done in [1], then it can be shown that for large γ , $H(t)$ can be approximated by the c.d.f $1 - e^{-\gamma \lambda^N t^N}$. a type 3 distribution of smallest values. From this it follows that there is an approximately linear relationship between the logarithm of the modal response time (or, more generally, the logarithm of any percentile of the response of the response time) and the logarithm of γ .

REFERENCES

- [1] Gart, J.J. (1965), "Some Stochastic Models Relating Time and Dosage in Response Curves," Biometrics 21 pp. 583-599.
- [2] Gumbel, E.J. (1954), "Statistical Theory of Extreme Values and Some Practical Applications," Nat. Bur. of Standards. Appl. Math. Ser., No. 33.
- [3] Mather, K. (1949), "The Analysis of Extinction Time Data in Biossay," Biometrics, 5, pp. 127-143.